EXHIBIT 15

S.E. CWIRLA ET AL

PROC. NATL. ACAD. SCI. USA 87 6378-6382, 1990

THIS PAGE BLANK (USPTO)

EXHIBIT 15

Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 6378–6382, August 1990
Biochemistry

Peptides on phage: A vast library of peptides for identifying ligands

(recombinant diversity/N-terminal hexapeptides/fd bacteriophage/avidity panning/antibody specificity)

STEVEN E. CWIRLA, ELIZABETH A. PETERS, RONALD W. BARRETT, AND WILLIAM J. DOWER*

Affymax Research Institute, 3180 Porter Drive, Palo Alto, CA 94304

Communicated by Lubert Stryer, June 6, 1990 (received for review May 3, 1990)

ABSTRACT We have constructed a vast library of peptides for finding compounds that bind to antibodies and other receptors. Millions of different hexapeptides were expressed at the N terminus of the adsorption protein (pIII) of fd phage. The vector fAFF1, derived from the tetracycline resistance-transducing vector fd-tet, allows cloning of oligonucleotides in a variety of locations in the 5′ region of gene III. A library of 3×10^8 recombinants was generated by cloning randomly synthesized oligonucleotides. The library was screened for high-avidity binding to a monoclonal antibody (3-E7) that is specific for the N terminus of β-endorphin (Tyr-Gly-Gly-Phe). Fifty-one clones selected by three rounds of the affinity purification technique called panning were sequenced and found to differ from previously known ligands for this antibody. The striking finding is that all 51 contained tyrosine as the N-terminal residue and that 48 contained glycine as the second residue. The binding affinities of six chemically synthesized hexapeptides from this set range from 0.35 μM (Tyr-Gly-Phe-Trp-Gly-Met) to 8.3 μM (Tyr-Ala-Gly-Phe-Ala-Gln), compared with 7.1 nM for a known high-affinity ligand (Tyr-Gly-Gly-Phe-Leu). These results show that ligands can be identified with no prior information concerning antibody specificity. Peptide libraries are also likely to be useful in finding ligands that bind to other classes of receptors and in discovering pharmacologic agents.

Biologically active molecules can be selected from large populations of randomly generated sequences. This approach has provided insight into protein targeting (1, 2) and enzymatic catalysis (3–5). A similar approach might be used to discover ligands for proteins and other receptors. Smith and Parmley showed that foreign DNA fragments can be inserted into gene III of filamentous phage to create infective "fusion phage" that display foreign peptides on their surface. The peptide inserts are accessible to antibodies, allowing purification of the phage and identification of the peptide sequences they carry (6, 7). They suggested that epitope libraries consisting of large numbers of clones expressing different short peptide sequences could be used to design vaccines, identify genes, and map epitopes (7). Subsequently, the screening of a small, model epitope library was reported (8). We describe here the construction of a phage library of 3×10^8 recombinants encoding millions of N-terminal hexapeptide sequences. This vast library was screened with a monoclonal antibody (mAb) specific for the Tyr-Gly-Gly-Phe sequence present in β-endorphin. The sequences of 51 high-avidity clones provide valuable information about the specificity of the antibody. Our study shows that randomly generated peptide sequences are a rich source of ligands.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Reagents and Strains. *Bst*XI restriction endonuclease, T4 DNA ligase, and T4 kinase were obtained from New England Biolabs; streptavidin and biotinylated goat anti-mouse IgG were from BRL; and Sequenase 2.0 was from United States Biochemical. mAb 3-E7, specific for β-endorphin, was provided by A. Herz and subsequently was purchased from Gramsch Laboratories (Schwabhausen, F.R.G.), and β-endorphin with [^{125}I]iodotyrosine replacing tyrosine at position 27 (β-[^{125}I-Tyr^{27}]endorphin; 2000 Ci/mmol; 1 Ci = 37 GBq) was from Amersham. Oligonucleotides were synthesized with an Applied Biosystems PCR-Mate and purified on OPC columns. Peptides were synthesized with an Applied Biosystems 431A synthesizer and purified to >95% purity by HPLC. Peptide content of the pure peptides was determined by amino acid analysis, and the composition was verified by mass spectroscopy. Bacteriophage fd-tet and *Escherichia coli* K91 were provided by G. Smith.

Construction of Vector fAFF1. A filamentous bacteriophage affinity vector, fAFF1, was constructed from the tetracycline resistance-transducing vector fd-tet (9). After removing the *Bst*XI site present in the transposon Tn*10* region of fd-tet, a cloning site consisting of two noncomplementary *Bst*XI sites was placed in the 5′ region of gene III. Site-directed mutagenesis was carried out with the oligonucleotide 5′-TAT GAG GTT TTG CCA GAC AAC TGG AAC AGT TTC AGC GGA GTG CCA GTA GAA TGG AAC AAC TAA AGG-3′ (10).

Construction of a Diverse Oligonucleotide Library. A collection of oligonucleotides encoding all possible hexapeptides was synthesized with the sequence 5′-C TCT CAC TCC $(NNK)_6$ GGC GGC ACT GTT GAA AGT TGT-3′ (ON-49), in which N is A, C, G, or T (equimolar), and K is G or T (equimolar). Two "half-site" oligonucleotides, ON-28 (5′-GGA GTG AGA GTA GA-3′) and ON-29 (5′-CTT TCA ACA GT-3′), complementary to the 5′ and 3′ ends of ON-49, were synthesized. The three oligonucleotides were phosphorylated with T4 kinase and annealed (in 20 mM Tris·HCl, pH 7.5/2 mM $MgCl_2$/50 mM NaCl) by mixing 1.5 μg of ON-28, 1.2 μg of ON-29, and 0.25 μg of ON-49, heating to 65°C for 5 min, and cooling slowly to room temperature. Vector fAFF1 replicative form DNA was digested to completion with *Bst*XI and precipitated with ethanol, and 20 μg was added to the annealed oligonucleotides to give a molar ratio of 1:5:100:100 for fAFF1/ON-49/ON-28/ON-29. These fragments were ligated by the addition of 20 units of T4 ligase and ATP to 1 mM and were incubated overnight at 15°C to produce gapped circular molecules. The ligated DNA was precipitated with ethanol, resuspended in water, and transformed by electroporation into *E. coli* MC1061. Five electrotransformations, each containing 80 μl of cells and 4 μg of

Abbreviations: mAb, monoclonal antibody; β-[^{125}I-Tyr^{27}]endorphin, β-endorphin with [^{125}I]iodotyrosine replacing tyrosine at position 27.

*To whom reprint requests should be addressed.

DNA, were performed by pulsing at 12.5 kV/cm for 5 msec (11). After 1 hr of nonselective outgrowth at 37°C in 2 ml of SOC medium (12), the transformations were pooled, and an aliquot was removed and plated on Luria–Bertoni (LB) agar plates containing tetracycline (20 μg/ml) to assess the transformation efficiency. The remainder was added to 1 liter of L broth containing tetracycline (20 μg/ml) and grown through approximately 10 doublings at 37°C to amplify the library.

Isolation of Phage. Phage from liquid cultures were obtained by clearing the supernatant twice by centrifugation (8000 rpm for 10 min in a JA10 rotor at 4°C), precipitating phage particles with polyethylene glycol (3.3% polyethylene glycol 8000/0.4 M NaCl), and recentrifuging as above. Phage pellets were resuspended in TBS (50 mM Tris·HCl, pH 7.5/150 mM NaCl) and stored at 4°C. In some cases, phage were isolated from plate stocks by scraping from the agar surface, resuspending in L broth, and purifying as described above.

Affinity Purification. Phage (10^{11} to 10^{12} infectious particles) were treated overnight with 1 μg of purified antibody in 1 ml of TBS at 4°C. Phage expressing peptides with affinity for mAb 3-E7 were isolated by a method known as panning, essentially as described by Parmley and Smith (7). Biotinylated goat anti-mouse IgG was added to the first-antibody-treated phage and incubated for 2 hr at 4°C. This mixture was then panned for 60 min on a plate coated with streptavidin. The phage solution was removed, and the plate was washed 10 times with 10 ml of TBS/0.05% Tween 20 over a period of 60 min at room temperature. Adherent phage were eluted by adding 800 μl of elution buffer [0.1 M HCl (pH adjusted to 2.2 with glycine) containing bovine serum albumin at 1 mg/ml] and incubating for 10 min at room temperature to dissociate the immunocomplexes; the eluate was removed and neutralized with 45 μl of 2 M Tris base as described (7).

Eluted phage were amplified by infecting logarithmic phase *E. coli* K91 cells, plating on LB agar plates with tetracycline, and growing overnight at 37°C. Phage were isolated from these plates, and the affinity purification process was repeated twice. After each round of panning and amplification, DNA of phage from several thousand colonies was pooled and sequenced. After the final round, a portion of the eluate was used to infect cells that were plated at low density on LB agar plates containing tetracycline. Individual colonies were picked, and phage DNA was isolated by a method designed for the Beckman Biomek Workstation using 96-well microtiter plates (13). The DNA was sequenced from a primer (5'-CGA TCT AAA GTT TTG TCG TCT-3') complementary to the sequence 40 nucleotides to the 3' side of the second *Bst*XI site in fAFF1.

Estimation of Binding Affinities of Peptides. Solution RIA using β-[^{125}I-Tyr^{27}]endorphin (20,000 cpm) and purified 3-E7 mAb (0.25 μg/ml) was conducted as described (14) except that the final volume was 150 μl. Antibody-bound and free β-[^{125}I-Tyr^{27}]endorphin were separated by addition of activated charcoal followed by centrifugation as described (15). Antibody-bound β-[^{125}I-Tyr^{27}]endorphin in the supernatant was measured in a γ counter. For each peptide, inhibition of β-[^{125}I-Tyr^{27}]endorphin was determined at six different concentrations at one-third logarithmic unit intervals, and the IC_{50} was determined by fitting the data to a two-parameter logistic equation with the ALLFIT program (16). Under the conditions of the RIA, the IC_{50} should be close to the dissociation constant (K_d) for the peptide.

RESULTS

Design of a High-Efficiency Oligonucleotide-Cloning Strategy. The vector fAFF1 was designed to provide many choices in the size and location of peptides fused to pIII. The two *Bst*XI sites flank the region encoding amino acids surrounding the signal peptidase cleavage site (the N terminus of the mature pIII) (Fig. 1). fAFF1 also contains a −1 frameshift mutation in pIII that results in noninfective phage. By removing the 31-base pair (bp) *Bst*XI fragment and inserting an appropriate oligonucleotide, (*i*) portions of the removed sequence can be precisely reconstructed (the correct signal peptide, for example), (*ii*) one or more additional amino acids may be expressed at several locations, and (*iii*) the correct translation frame in pIII is restored to produce infective phage.

Oligonucleotides to be cloned into fAFF1 have the general structure shown in Fig. 1A. The 5' and 3' ends have a fixed sequence, usually chosen to reconstruct the amino acid sequence in the vicinity of the peptidase cleavage site. The central portion contains one or more variable regions and also may code for spacer residues on either or both sides of the variable sequence.

"Half-site" oligonucleotides are hybridized to the 5' and 3' ends of oligonucleotide ON-49 to form appropriate *Bst*XI cohesive ends without the need to digest with *Bst*XI. This avoids cutting any *Bst*XI sites that may occur in the variable region. The *Bst*XI ends were chosen to provide oriented cloning of the oligonucleotides. The hybridized structure is then ligated to *Bst*XI-cut fAFF1 replicative form DNA to produce a double-stranded circular molecule with a small, single-stranded gap (Fig. 1*B*). These molecules are transformed by electroporation into host cells.

Construction of a Library of Hexapeptides Expressed on the N Terminus of pIII. The adsorption protein, pIII, is made as a precursor protein with an 18-amino acid leader sequence that directs pIII to the inner membrane of the bacterial host cell before assembly into intact phage particles (17, 18). We constructed a peptide library by cloning oligonucleotides of the structure shown in Fig. 1 to place the variable hexapeptide region at the N terminus of the processed protein. These first six residues are followed by two glycine residues (as a flexible spacer) and then the normal sequence of pIII. The library consists of 3×10^8 independent recombinants recovered as tetracycline-resistant colonies; 72% of these produced infective phage.

Phage DNA from a pool of several thousand of these colonies was sequenced to estimate the diversity in the cloning site. In the first two positions of each codon, bands of about the same intensity appeared in each lane, indicating the expected distribution of bases in these positions. In the

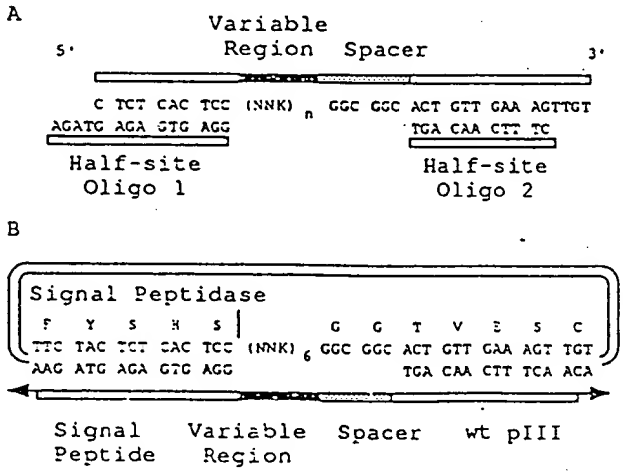

FIG. 1. Construction of the oligonucleotide library. (*A*) The oligonucleotide ON-49 is hybridized to two "half-site" fragments to form cohesive termini complementary to *Bst*XI sites 1 and 2 in the vector. (*B*) The vector fAFF1 contains two noncomplementary *Bst*XI sites separated by a 31-bp fragment. Removal of the *Bst*XI fragment allows oriented ligation of oligonucleotides with the appropriate cohesive ends.

third position of each codon, the G band was somewhat more intense than the T band.

To analyze the diversity of peptide sequences in the library in a more direct way, we picked 52 individual colonies producing infectious phage and sequenced the DNA of their variable regions. The distribution of bases at each position within the codons is given in Table 1. The first two positions have close to the expected equimolar distribution of the four bases. The third position contains about 50% more G than T in this sample. This bias is most likely introduced during the chemical synthesis of the oligonucleotide mixture but may also reflect biological selection. The amino acid sequences of the variable peptides of these 52 phage are given in Fig. 2. The ratio of the observed occurrence of each amino acid to that expected on the basis of codon frequency ranges from about 0.5 to 2 (Table 2), consistent with a random distribution of sequences.

Isolation and Sequencing of Phage Having High Avidity for Anti-β-Endorphin Antibody. mAb 3-E7 was raised against β-endorphin and, like the δ-opioid receptor, recognizes the N-terminal portion of the protein. The antibody also binds tightly to [Leu^5]- and [Met^5]enkephalin [Tyr-Gly-Gly-Phe-(Leu or Met)] and a variety of related opioid peptides (14, 19, 20). We have screened the N-terminal hexapeptide library against 3-E7 by carrying out three rounds of panning, elution, and amplification. This process selected 0.0048%, 0.25%, and 31% of the input phage in rounds 1, 2, and 3, respectively. The recovery of about a 100-fold greater proportion of the input phage with each succeeding round indicated that enrichment of certain phage was occurring.

After each round of panning, DNA representing several thousand eluted phage was pooled and sequenced. The area of the sequencing gel corresponding to the insertion site in gene III is shown in Fig. 3. The codon TCC specifying the serine that precedes the variable region is indicated by an arrow. After the first round of panning, the codon following this serine is clearly enriched in TAT (the single codon for tyrosine). After the second round, virtually all first codons in the pooled DNA appear to be TAT. The second codons are strongly GGK [the two codons for glycine; K = G or T]. Phage obtained after three rounds of panning appear to contain relatively few kinds of amino acids in the first four positions, whereas the fifth and sixth positions appear nearly as diverse as those of unselected phage.

The DNA samples from 51 individual phage recovered from the third panning were sequenced, and the deduced peptide sequences are shown in Fig. 4. The striking finding is that all 51 tightly bound phage analyzed had an N-terminal tyrosine. Furthermore, nearly all (94%) had glycine in the second position (Table 3). The third position is occupied by many amino acids, a few of which (e.g., glycine) are present more often than would be expected by chance. The fourth position is occupied primarily by the large aromatic residues tryptophan and phenylalanine (together 50%), and the bulky hydrophobic residues leucine and isoleucine (an additional 43%). The fifth and sixth positions contain essentially random distributions of amino acids.

Table 1. Nucleotide distribution in the variable region of infectious phage randomly selected from the library

Base	Frequency of each base by position in codon, %		
	1	2	3
G	31	27	59
A	22	22	<1
T	25	26	39
C	22	24	1

G M L Q R L	C R G N S G	D W V G G A	D L S P K V
M S R K L F	E R K A S V	T A M Q P G	D S E V S L
L W A G H E	W A E V F M	L D L K R L	R A A R D C
V G I T Q L	V A A G L N	V R N S M G	I Y T L H R
K T S Y G G	G P L P L F	R V C N K T	I T A P Y S
Q K R G E D	S N K G W A	R W S W E Q	S N D L S G
T L T K R Q	K H M L R W	G N M A H F	R S L H A G
S V S L Q A	V Q R L G K	F D S F G R	R W T W L G
A G S F E A	S G L Q R G	Q A V L M Q	A T L G F S
A I A A R A	W E K P R R	G K H Y Q W	I P G L L L
K A R L G L	R L V S T H	Q S R K S F	V C L L T V
L A F L A M	C A S L R S	G Y S S V D	G G G F T M
E R C R V D	Y A P S T R	S I G Q S K	V C P Q F C

FIG. 2. Amino acid sequences (deduced from DNA sequence) of the N-terminal hexapeptides on pIII of infectious phage randomly selected from the library. Sequences begin at the signal peptidase cleavage site. The single letter code for amino acids is: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

Binding Affinities of Peptides for mAb 3-E7. A solution RIA was used to estimate the binding affinities of peptides for mAb 3-E7 (Table 4). As previously reported (14), the 3-E7 mAb recognizes the N-terminal epitope Tyr-Gly-Gly-Phe, which is common to naturally occurring opioid peptides (14). Removal of tyrosine or deletion of any residues of the sequence Tyr-Gly-Gly-Phe-Leu drastically lowered the binding affinity (Table 4). Six hexapeptides with sequences corresponding to those of antibody-selected phage were chemically synthesized and assayed for binding to 3-E7. Their binding affinities ranged from 0.35 μM to 8.3 μM, orders of magnitude weaker than that of Tyr-Gly-Gly-Phe-Leu ([Leu^5]enkephalin) (7.1 nM) (Table 4).

DISCUSSION

We have generated a highly diverse peptide library and have demonstrated that specific ligands can be isolated by an approach that does not require prior structural information. The filamentous phage vector fAFF1 permits expression of peptides in a variety of positions in the N-terminal region of the phage adsorption protein, pIII. This vector was used to generate an N-terminal hexapeptide library of 3×10^8 independent clones.

Table 2. Amino acid content in the variable peptide of 52 randomly selected infectious phage

Amino acid	Nominal frequency	Occurrence		
		Nominal	Observed	Obs/Nom
A	0.065	19	27	1.42
C	0.032	9	8	0.89
D	0.032	9	10	1.11
E	0.032	9	9	1.00
F	0.032	9	12	1.33
G	0.065	19	33	1.74
H	0.032	9	7	0.78
I	0.032	9	6	0.67
K	0.032	9	16	1.78
L	0.097	28	35	1.25
M	0.032	9	10	1.11
N	0.032	9	7	0.78
P	0.065	19	9	0.47
Q	0.032	9	15	1.67
R	0.097	28	29	1.04
S	0.097	28	30	1.07
T	0.065	19	14	0.74
V	0.065	19	18	0.95
W	0.032	9	11	1.22
Y	0.032	9	6	0.67

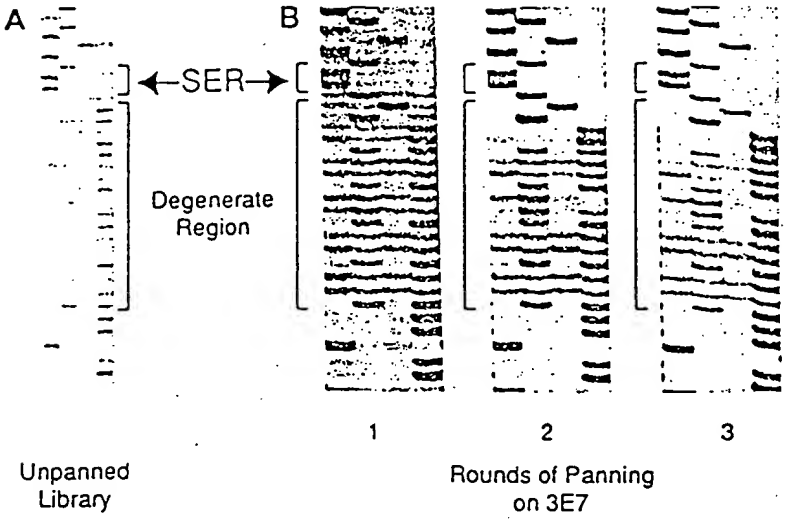

FIG. 3. Composite DNA sequence of the variable region of pools of infectious phage from the library (A) and phage recovered from one, two, or three rounds of panning on mAb 3-E7 (B). Phage were amplified as tetracycline-resistant colonies, and DNA from phage derived from a pool of several thousand of these colonies was isolated and sequenced. The area of the sequencing gel corresponding to the cloning site in gene III is displayed. A sequencing primer was annealed to the phage DNA ≈40 bases to the 3' side of the cloning site. The actual readout of the gel is the sequence complementary to the coding strand. For clarity of codon identification, the lanes may be read left to right as C, T, A, and G and top to bottom as 5' to 3' to identify the sequence of the coding (+)-strand.

The usefulness of the library as a source of ligands was assessed by screening it against a mAb with a well-known specificity. mAb 3-E7 binds the common N-terminal sequence Tyr-Gly-Gly-Phe present on all natural opioid peptides (14). In panning the library on 3-E7, we isolated a collection of phage with striking similarity; nearly all of the 51 phage we sequenced displayed Tyr-Gly on the N terminus of the variable peptide, a region known to be critically important for recognition by the antibody. A wider range of amino acids was present in the third and fourth positions, some at frequencies significantly greater than expected by chance. It is remarkable that no unrelated peptides were found after three rounds of panning, given the enormous diversity of sequences in the input phage population.

Scott and Smith (21) have generated libraries of 2.3×10^7 and 2×10^8 phage displaying diverse hexapeptides as internal epitopes four amino acids from the N terminus of pIII. They screened the smaller library by panning against antibodies specific for myohemerythrin and recovered two groups of phage, one bearing peptides with marked similarity to the natural epitope (21).

None of the peptides we identified by panning on mAb 3-E7 has been described previously as a ligand for this antibody. Six of these peptides were chemically synthesized and their binding affinities were estimated. Their apparent K_d ranged from 0.35 to 8.3 μM, significantly weaker affinities than that of the opioid peptide [Leu^5]enkephalin (Tyr-Gly-Gly-Phe-Leu) (K_d = 7.1 nM). It is possible that phage bearing peptides

YGGLGL	YGSLVL	YGALGG	YGWWGL	YGLWQS
YGGLGI	YGSLVQ	YGALSW	YGWWLT	YGFWGM
YGGLGR	YGSLVR	YGALDT	YGWLAT	YGXWSG
YGGLNV	YGSLAD	YGALEL	YGWANK	YGPFWS
YGGLRA	YGSLLS			YGEFVL
YGGLEM	YGSLNG	YGAIGF	YGNWTY	YGDFAF
	YGSLYE		YGNFAD	
YGGIAS		YGAWTR	YGNFPA	YAWGWG
YGGIAV	YGSWAS*			YAGFAQ
YGGIRP	YGSWAS*		YGTFIL	
	YGSWQA		YGTWST	YSMFKE
YGGWAG				
YGGWGP	YGSFLH		YGVWAS	
YGGWSS			YGVWWR	
YGGMKV				
YGGFPD				

*Identical nucleotide sequences.

FIG. 4. Amino acid sequences (deduced from DNA sequence) of the N-terminal peptides of pIII of 51 phage isolated by three rounds of panning on mAb 3E7.

with high affinities remained bound under the elution conditions, although this would require that the peptide–antibody interactions survive treatment at pH 2.2. These data indicate that the panning method is highly specific (no unrelated peptides were selected) but does not discriminate between those of moderate (micromolar K_d) and high (nanomolar K_d) affinities. The selection of phage-bearing peptides with relatively low affinity is likely the result of multivalent interactions between phage [carrying ≈4 copies of pIII (22)] and IgG molecules (containing two combining sites). Even with low-affinity interactions at each site, multivalent binding leads to high avidity and tenacious adherence of the phage during washing. In view of this avidity effect, it is not surprising that the few peptides sampled had moderate affinity. Among the 51 phage there was one pair with identical peptide sequences (these were also identical at the nucleotide level). A simple statistical calculation implies that about 2000 different peptides were probably present after the final round of panning. Moderate-affinity peptides are likely to be more abundant than high-affinity peptides, and the six clones we analyzed would most likely represent members of that lower affinity group. To enrich for the highest affinity ligands, monovalent interaction between phage and immobilized binding sites is probably required. Preliminary experiments with Fab' fragments of mAb 3-E7 support this contention.

Table 3. Distribution of amino acids in the variable peptide of 51 phage selected by panning with anti-β-endorphin antibody

Residue position	Amino acid	Frequency: Nominal	Frequency: Observed	Enrichment* ratio
1	Y	.031	1.00	33
2	G	.062	0.94	16
	A, S			<1
3	G	.062	0.28	5
	W	.031	0.10	3
	S	.093	0.21	2
	A	.062	0.12	2
	N	.031	0.06	2
	D, E, F, K, L, M, P, T, V			<1
4	W	.031	0.31	10
	F	.031	0.19	6
	L	.093	0.35	4
	I	.031	0.08	3
	A, G, M			<1

*Observed frequency divided by nominal frequency.

peptides for 3-E7 antibody determined by solution RIA

Peptide	K_d, μM	Peptide	K_d, μM
YGGFL	0.0071	YGFWGM	0.35
YGGF	0.19	YGPFWS	1.9
YGGL	3.8	YGGFPD	2.3
[illegible]FL	28	YGGWAG	7.8
[illegible]GG	>1000	YGNWTY	7.8
GGFL	>1000	YAGFAQ	8.3
GGF	>1000		
GFL	>1000		

Each K_d value is the antilogarithm of the mean of log K_d from three to six independent determinations with standard deviations of <0.25 logarithm units in all cases.

A strategy using a combination of conditions favoring multivalent and monovalent interactions may be used to advantage. By conducting the first rounds of panning under conditions promoting multivalent interactions, high-stringency washing can greatly reduce the background of nonspecifically bound phage. High-avidity binding selects a large pool of peptides with a wide range of affinities; and it may reveal specific recognition kernels, such as Tyr-Gly in the example described here. Subsequent panning on the basis of monovalent interactions enriches for the higher affinity peptides.

The oligonucleotides serving as the source of diversity in the library are synthesized with the degenerate codon motif NNK where K is G or T. This produces 32 codons, 1 for each of 12 amino acids, 2 for each of 5 amino acids, 3 for each of 3 amino acids, and 1 (amber) stop codon. Although this scheme provides the most equitable codon distribution available with standard methods of oligonucleotide synthesis, it results in a large bias against peptides containing 1-codon residues. For example, a peptide containing only 3-codon residues (e.g., hexaserine) is present 729 (3^6) times as often as a peptide containing 1-codon residues (e.g., hexaglutamine). Calculations based on this codon distribution and on the size of our library predict that about 75% of the 64 million hexapeptides would enter the library. A library of this size built with just 20 codons, each representing 1 amino acid, would initially contain about 97% of the possible hexapeptides. It is clearly preferable to generate random sequences by assembling 20 kinds of trinucleotides (one for each amino acid) rather than building from mononucleotides. This approach is especially attractive in generating longer peptide sequences, since the range of bias produced by the NNK structure increases 3-fold with each additional residue.

Biases in the distribution of peptides in the library may be caused by biological selection against certain of the peptides and by structural constraints on their availability for recognition. Constructing a library of peptides displayed on the N terminus of processed pIII necessarily alters amino acids in the vicinity of the signal peptidase cleavage site. Certain changes in the corresponding region of the major coat protein, pVIII, have been shown to reduce processing efficiency, slowing or preventing the incorporation of pVIII into virions (23). If pIII was affected similarly, the diversity of peptides contained in the library would be reduced. Our finding that most amino acids appear at each position of the variable peptides of randomly selected phage indicates that processing defects do not impose severe constraints on the diversity of the library.

Libraries of peptides on phage will undoubtedly be useful for mapping antibody epitopes. A much larger number of potential epitopes can be scanned than is feasible with methods based on chemical synthesis of candidate peptides (24). In addition, these libraries should be useful in discovering ligands for other important binding proteins such as hormone receptors and enzymes. For these purposes, libraries with different configurations may be constructed. In some cases, peptides longer than six amino acids may be required. Spacers of flexible or rigid structure may be used to vary the presentation of the variable peptides. Discontinuous regions of diversity, separated by one or more spacers, could serve as probes to locate multiple points of interaction in a large binding site. Although, in general, the peptide library screening approach does not require *a priori* information on ligand structure, such information when it does exist may be put to good use. The variable peptide can be arranged around or adjacent to a known recognition kernel [e.g., Arg-Gly-Asp in searching for ligands that bind to integrins (25)]. Future development and screening of peptide libraries should lead to the discovery of novel ligands for many purposes, including the identification of new drug candidates.

We especially thank Peter Schultz for suggesting the construction of a recombinant peptide library as a general source of easily identified ligands and for his many helpful suggestions during the course of this work. We also thank Lubert Stryer for his insights concerning the distribution of sequences in the library and for his comments, criticisms, and support. We appreciate the guidance of George Smith on the panning protocols and his gift of bacterial and phage strains. Avram Goldstein, Jeff Miller, and Steve Fodor made many useful suggestions concerning the experimental approaches. Jean Lee and Paul Yu Yang provided the chemically synthesized peptides.

1. Kaiser, C. A., Preuss, D., Grisafi, P. & Botstein, D. (1987) *Science* 235, 312–317.
2. Lemire, B. D., Frankhauser, C., Baker, A. & Schatz, G. (1989) *J. Biol. Chem.* 264, 20206–20215.
3. Dube, D. K. & Loeb, L. A. (1989) *Biochemistry* 28, 5704–5707.
4. Oliphant, A. R. & Struhl, K. (1989) *Proc. Natl. Acad. Sci. USA* 86, 9094–9098.
5. Hermes, J. D., Blacklow, S. C. & Knowles, J. R. (1990) *Proc. Natl. Acad. Sci. USA* 87, 696–700.
6. Smith, G. P. (1985) *Science* 228, 1315–1317.
7. Parmley, S. F. & Smith, G. P. (1988) *Gene* 73, 305–318.
8. Scott, J., Smith, G. P. & Dower, W. J. (1990) in *Advances in Gene Technology: The Molecular Biology of Immune Diseases and the Immune Response*, ed. Streilein, J. W. (IRL, Oxford), p. 224.
9. Zacher, A. N., Stock, C. A., Golden, J. W. & Smith, G. P. (1980) *Gene* 9, 127–140.
10. Kunkel, T. A., Roberts, J. D. & Zakour, R. A. (1987) *Methods. Enzymol.* 154, 367–382.
11. Dower, W. J., Miller, J. F. & Ragsdale, C. W. (1988) *Nucleic Acids Res.* 16, 6127–6145.
12. Hanahan, D. (1983) *J. Mol. Biol.* 166, 557–580.
13. Mardis, E. & Roe, B. (1989) *Biotechniques* 7, 840–850.
14. Meo, T., Gramsch, C., Inan, R., Hollt, V., Weber, E., Herz, A. & Riethmuller, G. (1983) *Proc. Natl. Acad. Sci. USA* 80, 4084–4088.
15. Ghazarossian, V. E., Chavkin, C. & Goldstein, A. (1980) *Life Sci.* 27, 75–86.
16. DeLean, A., Munson, P. J. & Rodbard, D. (1978) *Am. J. Physiol.* 235, E97–E102.
17. Shaller, H., Beck, E. & Takanami, M. (1978) in *The Single-Stranded DNA Phages*, eds. Denhardt, D. T., Dressler, D. & Ray, D. S. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 139–163.
18. Boeke, J. D. & Model, P. (1982) *Proc. Natl. Acad. Sci. USA* 79, 5200–5204.
19. Herz, A., Gramsch, C., Hollt, V., Meo, T. & Riethmuller, G. (1982) *Life Sci.* 31, 1721–1724.
20. Gramsch, C., Meo, T., Riethmuller, G. & Herz, A. (1983) *J. Neurochem.* 40, 1220–1226.
21. Scott, J. K. & Smith, G. P. (1990) *Science*, in press.
22. Goldsmith, M. E. & Konigsberg, W. H. (1977) *Biochemistry* 16, 2686–2694.
23. Russel, M. & Model, P. (1981) *Proc. Natl. Acad. Sci. USA* 78, 1717–1721.
24. Geysen, H. M., Rodda, S. J., Mason, T. J., Tribbick, G. & Schoofs, P. G. (1987) *J. Immunol. Methods* 102, 259–274.
25. Ruoslahti, E. & Pierschbacher, M. D. (1987) *Science* 238, 491–497.

THIS PAGE BLANK (USPTO)